Leveraging of Open EMR Architecture for Clinical Trial Accrual

Lawrence B. Afrin, M.D.[†], James C. Oates, M.D.[‡], Caroline K. Boyd, B.S.[§], and Mark S. Daniels, M.S.[§]
[†]Division of Hematology/Oncology, [‡]Division of Rheumatology, and [§]Center for Computing and Information Technology
Medical University of South Carolina (MUSC), Charleston, South Carolina

Abstract. Accrual to clinical trials is a major bottleneck in scientific progress in clinical medicine. Many methods for identifying potential subjects and improving accrual have been pursued; few have succeeded, and none have proven generally reproducible or scalable. We leveraged the open architecture of the core clinical data repository of our electronic medical record system to prototype a solution for this problem in a manner consistent with contemporary regulations and research ethics. We piloted the solution with a local investigator-initiated trial for which candidate identification was expected to be difficult. Key results in the eleven months of experience to date include automated screening of 7,296,708 lab results from 69,288 patients, detection of 1,768 screening tests of interest, identification of 70 potential candidates who met all further automated criteria, and accrual of three candidates to the trial. Hypotheses for this disappointing impact on accrual, and directions for future research, are discussed.

Introduction. Accrual to clinical trials has been and continues to be a major bottleneck in scientific progress in clinical medicine. In oncology, for example, fewer than 3% of potentially eligible patients enroll in trials.[1] This situation is particularly frustrating given that the current acceleration in biomedical discoveries is driving an increasing need for clinical trials.

Many methods for identifying potential subjects and improving accrual have been pursued. Most methods have focused on heightening the awareness of investigators, referring clinicians, and/or patients and the public. That commercial advertising has some effect is verified by the existence of the advertising industry, but the cost of sufficient commercial advertising is often prohibitive. Other methods of heightening awareness include paper and electronic flyers distributed by trial centers' internal mail systems, community and trial center bulletin board postings, contacts with patient support groups and advocacy organizations (e.g., the Susan G. Komen Breast Cancer Foundation[2] and the American Diabetes Association[3]), listings in trial registries (e.g., PDQ[4]), web sites (e.g., CenterWatch[5], clinicaltrials.gov[6], Yahoo clinical trials[7]), spam (mass e-mailings), and pocket computer-based trial databases and eligibility checkers.[8]

Although no controlled studies of methods of identifying potential subjects and improving accrual have been performed, it is generally acknowledged that few of the methods that have been employed have been appreciated by investigators to have had a significant impact, and none have proven generally reproducible or scalable, thus explaining why accrual remains the bottleneck described above.

Identifying potential subjects can be particularly frustrating in trials with especially stringent eligibility criteria or trials investigating uncommon diseases. In this regard use of the web for trial promotion appears to have had a significant impact on accrual for an occasional trial,[9] but the lack of general improvement in trial accrual to date despite the now widespread use of the web by the public[10-11] attests to the general lack of accrual impact by existing web-based trial promotion activity.

An alternative approach to identifying potential subjects is mass screenings. Where such screenings require human involvement (e.g., examination of patients by clinicians in the exam room or at a community event), resource limitations often decidedly constrain the "mass" part of "mass screening." However, often a trial's key eligibility criterion is a data element that has been recorded about a subject as a byproduct of an interaction with the subject totally unrelated to any trial activity. For example, a blood pressure routinely recorded at an annual check-up may identify the patient as a candidate for a hypertension trial.

Modern information systems make it theoretically possible to mass-screen any given data element at comparatively little cost, but in practice there have been challenges to such mass-screenings in the technical and ethical arenas, to which are now added regulatory challenges such as the privacy provisions of the Health Insurance Portability and Accountability Act of 1996 (HIPAA) which are now being implemented nationwide.[12]

Although a few vendors are now beginning to develop data warehousing and other functionality useful in clinical research, most clinical information systems today in use today are commercial offerings which provide no technical functionality for mass-screenings of specific data elements, and the database schemas and programming hooks that would be needed to locally develop mass-screening functionality are either unavailable or reserved out of proprietary interests by the vendors of most commercial clinical information systems.

There also are significant political/ethical considerations. Current clinical and research ethics clearly proscribe a third party making an investigator aware of a potential subject without the subject's advance consent for release of his identity to the investigator, and blanket consents (e.g., "I consent that any tissue or data obtained from me during my encounter may be used for research purposes and that our researchers are allowed access to my tissue and data") are generally seen as invalid, too.

We leveraged the open architecture of the core of our electronic medical record (EMR) system, and we worked with our institutional review board, to prototype a solution for these barriers to automated mass-screenings. There has been limited experience in the literature to date in automated mass-screenings of electronic clinical data for purposes of identifying potential trial subjects[13]. Our work adds to this experience.

Materials and Methods. MUSC uses the Oacis system (Dinmar, U.S., Inc., San Francisco, California) as its core clinical data repository and primary clinical data viewing application. MUSC's current implementation of Oacis include a "back end" Sybase database serving as the repository, which contains demographic and encounter data, diagnostic test results, provider notes, and other clinical data on more than half a million patients. The repository has been continuously collecting data since 1993 from an expanding array of best-of-breed enterprise and departmental systems, including the Cerner PathNet laboratory information system.

An MUSC rheumatologist (author JCO) opened a trial for lupus nephritis patients. Eligible patients are required to have certain suggestive diagnostic laboratory test results (e.g., significant proteinuria and a positive anti-nuclear antigen antibody or a positive anti-double-stranded-DNA antibody). Because of the open architecture nature of the repository and the interfaces into the repository, we were able to code a simplistic rule-based mechanism that could programmatically "watch" the Cerner-Oacis interface looking for results of a key screening lab test, thus identifying potential study candidates. This mechanism consisted of a Perl script scanning the messages flowing through the Cerner-Oacis interface. For every urinalysis (24-hour or random) in which an abnormally elevated protein level is found, a Sybase stored procedure is executed to identify whether the patient has ever had a non-negative ANA or anti-dsDNA test result. If so, the physician who ordered the urinalysis is notified of the patient's potential eligibility for the lupus nephritis study. All notifications are logged in a database for auditing purposes; physicians previously notified of a patient's potential eligibility are not re-notified.

Development (including coding and testing) of the Perl script and the Sybase stored procedure together cost approximately one week (40 hours) of time from an MUSC programmer familiar with the Oacis system.

In an MUSC Institutional Review Board-approved and HIPAA-compliant process specified in the trial's protocol, triggering of the rule generates notifications to the ordering physician advising him of the patient's potential eligibility and whom to contact for more information. (If the ordering physician's e-mail address cannot be identified, the notifications are sent to the attending physician.) The notifications consist of an e-mail message and a message to the physician's alphanumeric pager, which is serviced by the paging system owned and operated by MUSC. The IRB required the patient's identity be omitted from the pager messages but allowed inclusion of the patient's identity in the e-mail messages.

In order to identify the correct physician to notify, a linkage was needed between the ordering physician information found in the laboratory message and that physician's email address and pager number. Fortunately, as a result of MUSC's (still early) efforts to move toward "single sign-on" capability, an MUSC user's Oacis login name is the same as his MUSC e-mail account name. Furthermore, the physician number assigned by the patient billing system and associated with all orders (and subsequent results) is linked to the corresponding physician's Oacis login name. While e-mail notifications are sent at the time of criteria match discovery (24 hours a day), pager notification is limited to 8am-7pm. Audit logs are consulted prior to notification to eliminate duplicate notifications to the physician.

The principal investigator also receives e-mail and pager notifications that a potential subject identification

event has occurred, but the patient's identity is omitted in the interests of confidentiality.

It is left to the ordering physician to inform the patient of his potential eligibility, to solicit the patient's permission to contact the trial staff, and then to actually contact the trial staff.

Because of the small fraction of MUSC's several hundred attending physicians expected to be targeted by this experimental trial eligibility notification system, no education, training, or other advance notice regarding this experiment was provided to the MUSC physician community.

Results. In the ten months of experience with this prototype of an automated accrual assistance system, the Oacis repository received from the Cerner laboratory information system a total of 7,296,708 test results on 69,288 patients (across 800,500 encounters), including 1,768 results on the key screening test of relevance to the installed filter. After applying the additional criteria contained in the Sybase stored procedure, the filter triggered on 70 patients to date. Duplicate triggers on the same patient (resulting from repetitive lab testing) are suppressed. Of the 70 triggers, e-mail addresses for the ordering physician could not be found for 19 triggers. Of these 19, e-mail addresses could not be found for the attending physician for six triggers. Thus, 70 notifications were sent to the principal investigator, but only 64 notifications were sent to 30 distinct ordering (or attending) physicians. Of these 64 notifications, 52 were by e-mail only because the trigger fired between 7:00 p.m. and 8:00 a.m.

Records of post-notification actions unfortunately are incomplete, but it is clear that out of the 64 notifications to ordering or attending physicians, several caused the notified physician to take action. The principal investigator or the study coordinator have received about 1-2 follow-up inquiries per month from notified physicians. Most of these contacts have led to review of other study criteria and rapid identification of the patient's ineligibility. Occasionally, the physician has not remembered which patient the test was performed on, information which the principal investigator and study coordinator have not been provided by the system. In these cases, the physician was referred back to the e-mail notification he received, but the principal investigator and study coordinator have received no further follow-up in those cases.

Notifications were directly responsible for the accrual of three patients. In two cases, interestingly, the ordering physician also was the trial's principal investigator, who had not yet become aware of the patient's diagnostic laboratory test results at the time he received the filter's notification. Of these two cases, one notification was via both paging and e-mail, while the other occurred after 7:00 p.m. and therefore was via e-mail only. In the former case, the principal investigator immediately discussed the trial with the patient and successfully completed the recruitment of the patient to the trial during that same visit. In the latter case, the principal investigator promptly followed up with the patient, discussed the trial, and successfully completed the recruitment.

In the case of the third patient accrued to the study as a result of a notification, the ordering physician responded immediately to pager notification. Eligibility was confirmed by phone and the patient subsequently visited the principal investigator and was successfully recruited at that visit.

Discussion. The need for improved rates of potential subject identification and accrual is widely recognized. Because of the ambiguous nature of many clinical trial eligibility criteria and a variety of problems with the contents of many clinical data repositories, completely automated eligibility determination remains a distant target for most trials. However, an automated screening of selected clinical parameters prior to a full manual screening may be a useful approach toward improving accrual — as long as such screening complies with current clinical and research ethics and applicable regulations (e.g., the health information privacy provisions of HIPAA, which forbid nonconsensual release of patient information to a third party not involved with treatment, payment, or other routine operations associated with the provision of health care to the patient).

It is worth noting, too, that mass-screenings need not be confined to lab results. Any discretely recorded clinical data or events (e.g., vital signs, medication orders) could be used. Natural language processing (NLP) of free-text clinical reports perhaps could be used, too; NLP accuracy, and therefore its utility in identifying potential trial candidates, has been improving steadily over the years but remains significantly variable from one report to the next.[14]

There has been limited experience in the literature to date in automated mass-screenings of electronic clinical data for purposes of identifying potential trial subjects. In 1995 Carlson et al. reported on screening a selected group of 60 HIV patients for eligibility for 17 protocols over a seven-month study period, identifying 165 accrual opportunities in 13 patients.[13] Most of these opportunities led to the discovery of reasons for ineligibil-

ity, but a complete accounting for the failure to accrue even a single patient to a single protocol was not provided.

We developed an alternative approach, performing truly a mass-screening of all-comers, with notifications of screening "hits" provided via methods we felt fit well with the workflows of our institution's physicians. Our approach yielded both positive and negative results.

On the positive side, we demonstrated that an open architecture model of the clinical information systems involved in such screening permits a rapid technical implementation of the screening. The open architecture of the Oacis repository provided us easy access to the Cerner-Oacis interface as well as to additional data elements needed for the overall process to work (e.g., physicians' pager codes). Had the repository been of closed architecture, or had the repository and laboratory information system been merely separate modules within a monolithic system, insertion of the filter and the notification logic likely would have been considerably more challenging, perhaps even impossible without vendor involvement.

Also on the positive side, a large number of potential subjects were identified by the system. In author JCO's experience, a potential subject identification rate of six per month significantly exceeds what would be expected through traditional trial awareness promotional activities.

Of course, the key negative result is the poor follow-through seen on the parts of the ordering physicians who received the automated notifications, resulting in a somewhat disappointing impact this automated accrual assistance system has had on accrual to this trial to date.

The actual reason(s) for this poor follow-through are unclear. One possibility for the poor follow-up we observed is that the pager notification message was vague in that it omitted any patient identification information. Again, we were constrained from providing this information by IRB mandate, which may have been concerned about the potential for inappropriate interception of the signal broadcast by the pager system, even though the system itself is owned and operated by MUSC.

The identities of the notified physicians were available to us. After adequate time had elapsed for the notified physicians to follow-up on their notifications, those physicians who had not followed up were solicited by e-mail for an explanation. No responses to these solicitations were received. Thus, we cannot know their reasoning for not following up, but it seems logical to assume that these physicians perceived that the cost of the follow-up (primarily their time) exceeded the potential benefit (to themselves, to their patients, to their investigator-colleague, and to medical science).

If this assumption is true, then the value of this system might be improved by decreasing the cost and/or increasing the potential benefit.

If the primary cost of the follow-up is the time that would have to be spent by the ordering physician in further interactions with the patient regarding the trial, one potential solution to this problem is for the investigator to provide a web site containing information about the trial and for the system to provide in the e-mail and pager notification messages a reference to the web site which the ordering physician can provide to the patient.

Another solution would be to entirely eliminate the ordering physician's involvement and have the system directly send notification to the patient of his potential eligibility, inviting the patient to contact the trial staff and/or visit the trial's web site for more information. (Alternatively, the patient could be invited to contact the ordering physician, who, upon knowing of the patient's awareness of the trial, may be more inclined to provide a referral to the trial staff.) Notification to the patient could be sent via standard mail or e-mail, if the repository contains the appropriate addresses.

The potential benefit to the ordering physician perhaps can be increased in multiple manners, too. For example, the notification message perhaps could offer the physician a modest "finder's fee," which could take a variety of financial or non-financial forms, if the physician follows up and the patient is successfully accrued. Or perhaps the notification message needs to more clearly identify the potential benefit to the patient.

Refinements of the notification process such as have been discussed above will need to be investigated so that greater system utility can be demonstrated before it will be appropriate to consider making this system generally available to MUSC investigators. Also, the centralized rules engine in a soon forthcoming upgrade to the Oacis system should permit even more rapid development of the screening filters.

Given that (1) information systems increasingly must coexist and interact with other systems in ways the designers cannot anticipate, (2) the cost of software development is proportional to the technical difficulty of the project, and (3) development of an interface with a closed system is considerably more technically difficult

than with an open system, then it logically follows that the open architecture approach to system design becomes not merely convenient but in fact critical if a vendor wants to see its system perceived as a viable solution by the customer. Open architecture of the core repository is especially important in academic medical centers, where the imperatives for leading-edge medical research and development activities often require the ability to use repository data in ways a repository vendor may not be able to anticipate.

Finally, it is worth noting that the recent development of (1) the concept of web services, (2) the standards associated with this concept, and (3) XML models of many domains of knowledge and data, all portend a much greater degree of open architecture among systems of all types, including health care systems, in the near future. Dinmar, U.S., Inc., whose Oacis system long has been recognized as the industry leader of the open architecture movement in the health care information systems arena, began migrating the Oacis system to an XML- and web services-based model two years ago to further enhance the system's openness. It is encouraging to see many other vendors (e.g., Cerner) now beginning this migration as well.

In conclusion, we leveraged our clinical data repository's open architecture to devise a previously unreported, IRB-approved method of mass-screening clinical laboratory results to identify potential candidates for a clinical trial for which recruitment was anticipated to be difficult. We identified far more potential candidates than would have been expected without this method, and the recruitment of three patients was directly attributable to this method. Physician response to notifications of potential patient eligibility have been relatively disappointing, for largely unclear reasons. We are investigating refinements to our method prior to making it generally available to our institution's clinical research community.

References

1. Winn RJ. Obstacles to the Accrual of Patients to Clinical Trials in the Community Setting. *Semin Oncol* 1994;21(4 suppl 7):112-7.
2. Susan G. Komen Foundation, http://www.komen.org
3. American Diabetes Association, http://www.diabetes.org
4. Physicians Data Query (PDQ), http://www.nci.nih.gov/clinical_trials.
5. CenterWatch, http://www.centerwatch.com
6. ClinicalTrials.gov, http://clinicaltrials.gov
7. Yahoo! Directory of Clinical Trials Websites, http://dir.yahoo.com/Health/Medicine/Clinical_Trials/
8. Breitfeld PP, Weisburd M, Overhage JM, et al. Pilot study of a point-of-use decision support tool for cancer clinical trials eligibility. *J Am Med Inform Assoc* 1999 Nov-Dec;6(6):466-77.
9. Masys DR. Improving Patient Accrual to Clinical Research Protocols Using the World-Wide Web. *Proc Am Med Informatics Assoc* Spring Congress, June 1996.
10. Pew Charitable Trusts. Counting on the Internet: Project Report of the Pew Project on the Internet and American Life, December 29, 2002, http://www.pewinternet.org/reports/toc.asp?Report=80
11. Greenspan R. More Than Half-Billion Online Globally. cyberatlas.com, February 21, 2003, http://cyberatlas.internet.com/big_picture/geographics/article/0,,5911_1593591,00.html
12. U.S. Health Insurance Portability and Accountability Act of 1996, http://www.cms.hhs.gov/hipaa/
13. Carlson RW, Tu SW, Lane NM, et al. Computer based screening of patients with HIV/AIDS for clinical trial eligibility. *Online J Curr Clin Trials* Document Number 179, 1995 Mar 28.
14. Bates DW, Evans RS, Murff H, et al. Detecting Adverse Events Using Information Technology. J Am Med Inform Assoc 2003 Mar-Apr;10(2):115-28.